

Architecture/Design for IOC-1

K. Berkgbigler, B. Bush, J. Davis

Los Alamos National Laboratory

6 February 1995

The purpose of this review is to outline the architecture and high-level design for TRANSIMS.

Although it specifies the basic requirements for IOC-1, it also provides a framework into which to incorporate the future IOCs.

The basic goal is to make sure that the various components of TRANSIMS are efficiently integrated, both for IOC-1 and beyond, so that TRANSIMS remains flexible, expandable, and maintainable throughout its lifetime.

A document should be maintained to incorporate changes to the architecture and requirements as they arise.

Outline

Architectural Specifications

Data Flow Specifications

Review of Previous Prototypes

Assessment of Major Risks

Architectural Specifications

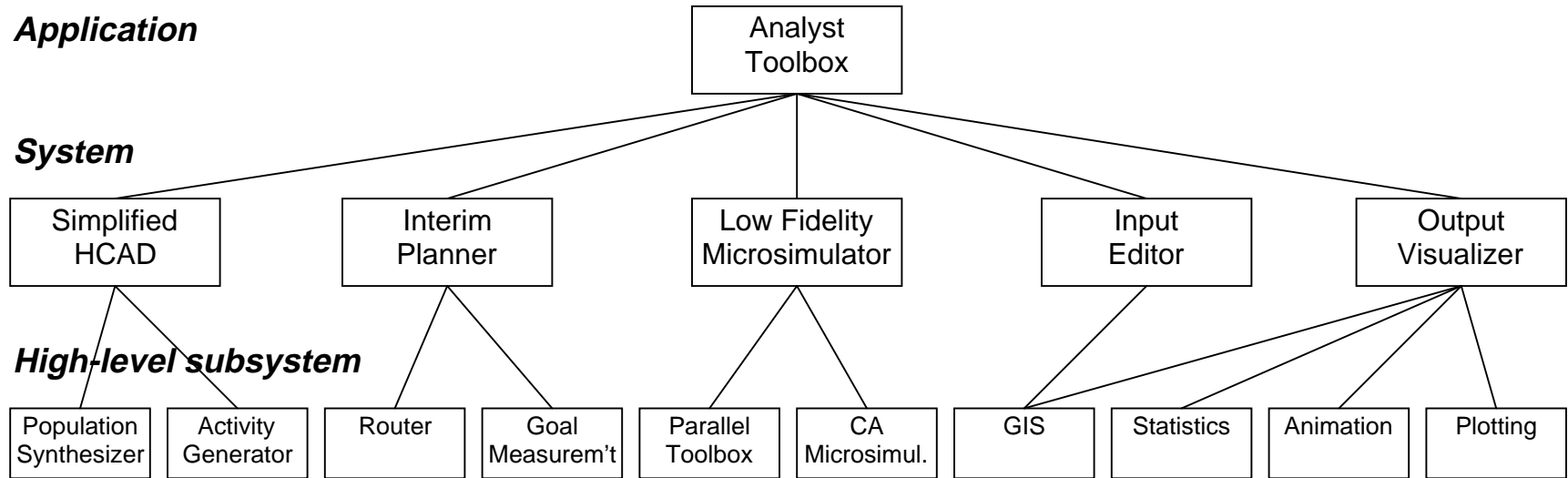
Layered Architecture

- Application: The Analyst Toolbox provides a centralized interface for TRANSIMS.
- System: The TRANSIMS systems (Input Editor, Simplified Planner, Low Fidelity Microsimulator, Output Visualizer) centralize access to the major functional components of TRANSIMS. Note that additional systems (such as a Disaggregator, High Fidelity Microsimulator, or Land Use Estimator) may be added for future IOCs or from sources external to LANL.
- Subsystem (high-level): The high-level subsystems each provide services to one or more TRANSIMS systems. This enhances the reusability and flexibility of the software. (Router subsystem could be used for planning by the Simplified Planner and for replanning by the Microsimulator.)

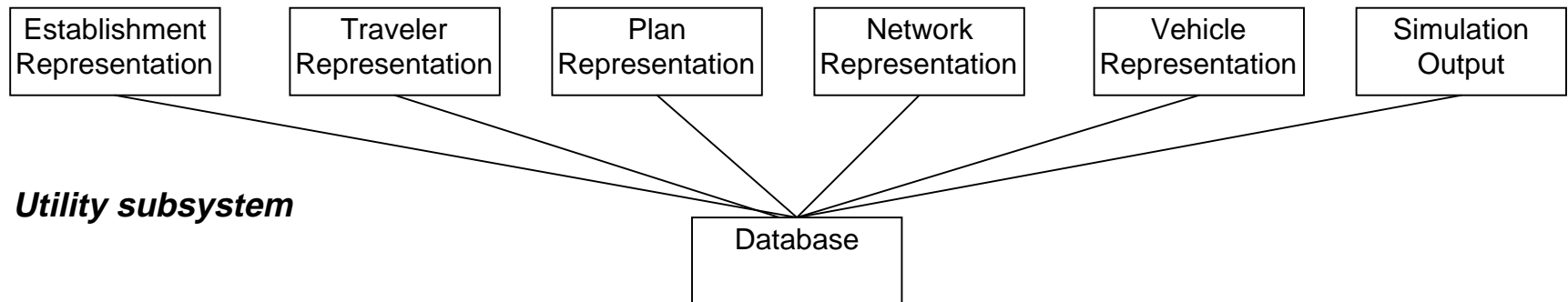
- Subsystem (low-level): The low-level subsystems provide basic services (mostly, data and operations on data) to the high-level subsystems and to the systems. They provide a common representation of objects such as vehicles, travelers, the road network, etc. They never directly interact with the user.
- Subsystem (utility): The utility subsystems provide basic domain-independent services to the higher-level components of TRANSIMS. They are used to isolate the domain subsystems from dependence on the operating system, file system, etc.

Application

System



Low-level subsystem



Utility subsystem

Analyst Toolbox

The analyst toolbox provides a centralized and uniform user interface integrating the major TRANSIMS systems.

System definitions/requirements

Input Editor

The input editor provides a means for editing the database and setting up scenarios for simulation.

All scenario data will be editable. It will have functions for:

- importing new road networks
- altering existing networks
- merging networks
- extracting networks
- editing traveler characteristics
- editing establishment characteristics
- editing vehicle characteristics

The editor will be integrated into the rest of the GIS software, supporting visual/graphical editing of geographic objects, table editing of non-geographic objects, and editing via ad-hoc queries (using the native query languages for Arc-View, Arc/Info, and Oracle).

Version control will also be supported.

Simplified Household and Commercial Activity Disaggregator (HCAD)

The simplified HCAD system provides the ability to construct travel activities for study areas.

This system will have the capability to select and import any of the available synthetic populations for a study region and to generate travel activity from the chosen synthetic population.

The focus in IOC-1 will be on households, as opposed to commercial activity.

Simplified Planner

The Simplified Planner will provide the microsimulator with household or individual travel demand in the form of trip plans.

For IOC-1, the major effort will be to develop the data structure which integrates the "flow" of data from the synthetic population and its activity demand, through the planner's trip planning process, and on to the microsimulation.

The planner's process, or mode/route/activity assignment algorithms, will be simplified to accommodate available data and microsimulation development time constraints.

There are three types of input data: the transportation network, household/individual activity demand, and individual travel behavior.

The network required by the planner is a subset of the total network representation.

Activity demand and travel behavior are derived from national and local demographic and socioeconomic characteristics, and local land use. For IOC-1, the Simplified HCAD will provide this data.

While the household is the basic unit of activity input, the emphasis in IOC-1 will be to identify mandatory activities at the individual level. Discretionary activities at the household or individual level may be included.

Individual travel behavior input will consist primarily of a limited set of goals and preference distributions.

The trip planning process will consist of activity, mode and route assignment algorithms. For IOC-1, no activity assignments will be made; the mandatory activities will be provided by the Simplified HCAD.

No modal assignments will be made because of the focus on the automobile. Carpooling will not be addressed; pedestrian movement may be addressed but only with respect to movement to and from the automobile.

Output data are the individual trip plans required by the microsimulation.

This data includes trips and trip chains consisting of origins, activity destinations, routes and times associated with activity performance and route movement.

Low Fidelity Microsimulator

The low fidelity microsimulator is a regional-scale, low-fidelity traffic microsimulation based on cellular automata and implemented on a distributed computer network. It will be the primary technical focus of IOC-1.

A preliminary design for this type of microsimulation was proposed by Marcus Rickert in "A First Draft on how to Integrate High Fidelity and Cellular Automata Approaches to Microsimulation in TRANSIMS on a Distributed Computer Network, Version 1.0," July 18, 1994.

The user interface to the low fidelity microsimulator will provide a means for specifying CPNs, load balancing parameters, diagnostic outputs, and simulation outputs, as well as data sources to be used as inputs.

Output Visualizer

The output visualization system provides an integrated interface to the various tools available for viewing and analyzing aspects of simulation and plan data. There are five types of output available:

- animation
- statistical analysis
- geographic analysis
- plots
- data export

The user can select a geographic region of interest (and time interval) and filter the available data before visualizing it. The output visualization system will retrieve the specified data from the data sources and then send the data to the appropriate visualization tool.

The output visualization system interface will handle as much of the user interaction as possible. Some of the interaction may have to take place in the third-party software that is integrated into the output visualizer.

The results of statistical and geographic analyses can also be plotted. It must also be possible to export data to standards formats for spreadsheets, etc.

Subsystem definitions/requirements

Representation Definition

A “representation” is a view of a database, plus utility classes for using the domain data.

The specification of a representation says nothing about the actual data organization. Rather, it specifies what data must be available.

Functions for both reading and writing data must be a part of a representation.

The representation subsystems will also contain utility functions for manipulating and using the data.

Network Representation

The transportation network representation includes detailed information about roads, intersections, signals, sensors, transit systems, rail, bikeways and walkways.

Network topology is represented along with attributes that describe the nodes and links in the network.

Multiple views of the network representation are required, in some degree, by the Planner, Microsimulation, and Output Visualizer.

Link attributes for the road network include such characteristics as link type, length, directionality, speed limit, number of lanes, special lane designators, grade, toll, passing allowed, visibility range, street name, traffic capacity, etc.

Node attributes may include node type, associated intersection (if applicable), etc. Provision will be made to represent the vehicular traffic network at a resolution at least as detailed as that of TRAF and TRANSYT-7F.

Intersections may be represented at multiple levels of fidelity. The Planner requires less fidelity than the Low Fidelity Microsimulator which uses a medium fidelity representation incorporating multiple queues. This intersection representation will include algorithms for handling the queues.

Signalized and unsignalized intersections will be represented, and the representation must include lane usage and allowable turning movements. Both timed and actuated signals will be included along with sensors for actuated signals. The signal representation will include algorithms for cycling through signal phases.

Transit networks, associated transit schedules, and inter-modal transfer facilities will not be represented in IOC-1, but provision will be made to include them in a later IOC.

The user interface to the network representation will be provided through the Input Editor. The network subsystem will support the importation of external data in formats such as Arch/Info, TRANSYT-7F, and the TRAF family.

Establishment Representation

The definition of an establishment includes households, group quarters, and businesses.

Each establishment will possess a unique identifier and the socioeconomic attributes required by the activity generator subsystem.

IOC-1 will not focus on commercial activity or the movement of freight. Provisions will be made for the needs of the more sophisticated planner and disaggregator to be developed in a later IOC

Traveler Representation

The traveler representation includes the demographic, socio-economic, and geographic attributes needed for identifying travelers and for planning trips.

It also includes the driver representation and driver model needed for simulating driver behavior.

The external data resources will not be available for the simplified planner. Thus the traveler representation for IOC-1 will be fairly simple.

Nevertheless, it should implicitly specify unique travelers and make it possible to associate them with vehicle identifiers in the plan representation.

The framework developed will be consistent with the future enhancements to be made in the planner and disaggregator.

The driver representation describes the attributes of the drivers used in the models. Driver's decision making processes may not be modeled at the same level of fidelity in every simulation, so a flexible driver representation that supports a particular decision logic without requiring unnecessary attributes is needed.

Driver attributes will be specific parameters that are required by the decision logic algorithms rather than abstract behavioral attributes such as aggressiveness.

Potential attributes include driver age, sex, socioeconomic attributes, desired speed, following distance, and acceptable gaps for left and right turns, crossing intersections, and changing lanes.

Algorithms that define the driver's decision logic are also part of the driver representation.

Traveler attributes required by the planner include household id, mandatory activities, age, sex, socioeconomic class, trip goal weights, and mode/route preference distributions.

Driver attributes included in the current CA Microsimulation include desired speed, and a variable that causes drivers to vary from their desired speed some of the time.

The user interface to the driver representation will support user specification of values for the attributes required in the study as well as the ability to read existing descriptions of driver properties.

Vehicle Representation

The vehicle representation describes the attributes of the vehicles used in the models.

Not all potential attributes are needed for every type of study, e.g. air quality studies require dynamic information about engine properties, while other studies may not. The vehicle representation should be flexible enough to support a variety of studies but not require unnecessary attributes for a particular study.

Potential attributes include such properties as vehicle type, maximum speed, maximum acceleration, maximum deceleration, stopping distance, length, width, weight, age, fuel type, and engine properties.

Dynamic attributes such as position along road segment, lane, velocity, acceleration, and engine temperature may also be required.

Algorithms that define the motion of vehicles are also part of the vehicle representation.

The version of the Planner to be used in IOC-1 requires only vehicle type. The current CA Microsimulation requires only maximum speed, but length will also be required if multiple vehicle types are supported. Acceleration and deceleration parameters may also be desirable to smooth speed fluctuations.

The user interface to the vehicle representation is provided through the Input Editor and will support user specification of values for the attributes required in the study as well as the ability to read existing descriptions of vehicle properties.

When multiple types of vehicles are modeled, fleet mix will also be under user control.

Plan Representation

The plan representation will provide a view of trips and trip chains in the form of routes, origins, and destinations during specific time periods.

A trip chain is a sequence of trips.

A trip must provide a:

- trip purpose
- unique identifier for the traveler
- unique identifier for a vehicle
- starting node in the network
- desired departure time from that starting node
- destination node in the network
- desired arrival time at that destination node

In the microsimulation, detailed routes will specify which streets simulated vehicles will follow and when they should be at various points along the way to satisfying an associated trip goal.

The detailed route design will impact the complexity of both the microsimulation and the visualization systems. The actual format might not include the ID of every link and node along the way. Perhaps a detailed route should be included as part of the trip specification.

Population Synthesizer

Using 1990 census data a baseline synthetic populations of households will be generated which statistically mimic those sampled in the 1990 census.

These populations will be produced on a census tract or block group basis.

Each household and person in the synthetic population will be associated the entire suite of socioeconomic characteristics available from the census.

For future applications these households will be aged to the desired date using projected land use and demographic trends in the study region. The populations will not be aged for IOC-1.

Input data for the generation of the baseline population will include the Census Bureau Standard Tape File 3 and the Public Use Microdata Sample.

Standard statistical techniques will be used to generate the synthetic populations.

For IOC-1 baseline synthetic populations will be generated off-line. Multiple populations will be produced, but the software for doing so is not part of IOC-1.

The households of these populations will be placed at random at locations in the census tract or block group.

Activity Generator

Travel activity will be predicted for each household and household member using national trends and local activity surveys.

These desired travel activities will be passed to the planner for routing and scheduling.

Activities will be assigned to either the household or the individual. For example work activities would be assigned to the individual while a household activity could be a shopping trip.

Additionally, activities will be either mandatory or discretionary.

The structure of IOC-1 will allow for future activity based travel. However, for the applications considered in IOC-1, travel activity will be replaced by trips which are generated from OD matrices.

All trips will be assigned to individuals (no household activity analysis is planned for IOC-1) and all trips will be treated as mandatory.

Trips from an OD matrix will be matched with the demographics for the households.

These will be randomly distributed to individuals from demographically matched households within the census tracts or block groups which make up the Dallas Travel Survey Zones.

OD Matrix / Route Disaggregator

For IOC-1, actual travel data is available only in the form of OD matrices and, possibly, traffic and turn counts.

To produce trips, a utility preprocessor will be needed to disaggregate OD matrix zonal based traffic flow down to individual travelers, specific routes and activity addresses.

Methods are currently being examined which will disaggregate the zonal flows along the boundaries of the detailed area of interest (e.g., along the LBJ corridor).

Once the individual travelers are placed on routes which enter the area of interest and assigned "start" times, the Simplified Planner will then perform the normal route assignment to generate detailed individual trip plans.

The travelers will move to final destinations within the area of interest or through and out of the area of interest based upon their aggregate OD matrix assignments.

If traffic or turn counts are available for the area of interest, they will be used to calibrate the planner's preference distributions.

Router

Given household or individual activity demand, individual travel behavior, and individual travelers along the boundaries of the detailed area of interest, this subsystem will provide feasible or "optimally" infeasible sets of trip plans.

A feasible trip plan is one in which all individual goals have been satisfied; an "optimally" infeasible plan is one which has not satisfied all of the individual's goals but which has minimized the non-zero goal deviations.

The router will be structurally consistent with the mode, route and activity assignment enhancements planned for later IOCs.

Parallel Toolbox

The parallel toolbox subsystem is those parts of the Low Fidelity Microsimulator that deal with running on a parallel distributed computer.

The parallel toolbox provides a master/slave parallel computing model implemented on top of the PVM toolkit for heterogeneous network computing. PVM provides the message-passing substrate that allows tasks on different machines (CPNs) to communicate.

Using the parallel toolbox, dynamic load balancing of the road network and associated vehicle objects is provided and is based on usage statistics gathered while the simulation runs.

CPNs may be added or deleted during a running simulation.

Diagnostic outputs will be provided as will some form of fault tolerance.

Simulation outputs will be produced in parallel.

CA Microsimulation

The CA microsimulation subsystem is those parts of the Low Fidelity Microsimulator that deal with doing traffic simulation using cellular automata.

The network representation in this approach is grid based, and a mapping from the general network representation subsystem to grids will be provided.

Similarly, mappings from the vehicle and traveler representation subsystems into vehicle/driver combinations that are suitable for the CA approach to vehicle motion will be made.

Travelers will utilize plans from the plan representation subsystem.

Simulation Output

The simulation output subsystem will gather the data generated by simulations and provide access to it for other subsystems as soon as the data is received.

The simulation output will be configurable and several predefined configurations will be provided:

- trajectory information (time, segment, lane, position along segment, velocity, acceleration)
- control systems (signals, sensors, HOV lanes)
- vehicle state information (brakes on, lights, signaling)
- measures of effectiveness (VHT, VMT, average speed, average density, headway)
- animation output (trajectory and control systems)
- engine performance (idle time, start time, stop time, temperature, fuel consumption)

- emissions (CO, NO_x, O₃, particulates, aerosols)
- traveler characteristics (vehicle occupancy, demographics, trip purpose, plan fulfillment)
- additional outputs available from TRAF and TRANSYT-7F products

This subsystem will utilize the database subsystem to support metadata, data distribution, data export, and archiving.

Special provision will be made for dealing with the large amount of data generated by simulations. It will also be possible to perform compression (lossy or lossless) on the data to reduce the storage required for it.

(Only limited capabilities will be developed in this area for IOC-1.)

In IOC-1, no automated feedback mechanism will be provided for using simulation output to alter simulation or planner input for subsequent simulations.

This can be accomplished, manually, through a (possibly complicated) series of ad-hoc queries involving the output and input data sets.

GIS

The GIS subsystem provides high-level support for the following functions:

- editing network data
- viewing all types of geographic data and non-geographic data that can be linked to geographic data
- thematic display and analysis of plans and simulation output
- aggregation/disaggregation of geographic data
- preprocessing/formatting of geographic data from external sources (i.e., MPO data)

The GIS subsystem will support the import/export and editing of data available from other data-related subsystems (database, network, traveler, vehicle, plan, output) as well as support the export of data to visualization and analysis tools such as the plotting, statistics, and animation subsystems.

Statistical Analysis

This subsystem will support the following general types of statistical analyses:

- confidence intervals
- analysis of variance
- hypothesis testing

All of the data available from other data-related subsystems (database, network, traveler, vehicle, plan, output, GIS) will be importable for analyses.

Predefined and user-definable analyses will be available for computing various measures of effectiveness; it will also be possible to save an analysis configuration for later recall and re-running.

The results of analyses will be exportable to the plotting subsystem.

Animation

This subsystem will provide animated display of vehicle movement and traffic controls (e.g., signals) in real time and in accelerated time.

Other map features such as buildings, bodies of water, topography will not be displayable for IOC-1.

The vehicles and network features will be selectable with the mouse to obtain detailed attribute data. It will be possible to color-code vehicles based on their attributes.

Plans can also be animated.

Plotting

This subsystem will support the following general types of plots:

- scatter plot
- histogram
- lines
- areas

Three-dimensional and color plotting will be supported, as will be the grouping of attributes.

All of the data available from other data-related subsystems (database, network, traveler, vehicle, plan, output, GIS, statistics) will be importable for plotting.

Predefined and user-definable plots (whose configuration can be saved and loaded) will be available.

It will also be possible to customize axes, legends, markers, and annotation.

The specific types of plots will include the following:

- waterfall plots
- fundamental diagrams (all styles)

Database

The database subsystem provides low-level services for accessing and modifying data.

It forms a layer separating the other subsystems from the actual data files—the other subsystems will not have access to the data files at the physical level.

Each data source will be indexed by a unique (primary) index. Additional (secondary) indexes will be allowed.

The interaction between other subsystems and the data will be mediated by the public interfaces of the classes in the database subsystem.

Range lookup of key values and ad-hoc SQL queries will be supported.

Procedures (e.g., C++ templates, preprocessor macros, or custom preprocessor) will be available to coordinate the maintenance of the database schema and class definitions.

The database subsystem will maintain metadata specifying the following for all of the data sources:

- existence
- versions
- network location
- attributes (data dictionary)

The distribution of a data source over the network will not be supported for IOC-1, although this capability will be added in a later IOC.

Migration of data will also be supported, but the database subsystem will not be required to automatically load-balance the data among the network nodes.

An archiving facility will also be available.

There will also be a mechanism for extracting data specified at run-time from a data source and exporting it to a formatted binary or ascii file.

System-Subsystem Relationships

<i>Subsystem</i>	<i>Input Editor</i>	<i>Simple HCAD</i>	<i>Simple Planner</i>	<i>Low Fidelity Sim.</i>	<i>Output Visual.</i>
Network Representation	X		X	X	X
Establishment Represent.	X	X	X		X
Traveler Representation	X	X	X	X	X
Vehicle Representation	X		X	X	X
Plan Representation	X	X	X	X	X
Population Synthesizer		X			
Activity Generator		X			
OD Matrix/Route Disag.			X		
Router			X		
Parallel Toolbox				X	
CA Microsimulation				X	
Simulation Output				X	X

<i>Subsystem</i>	<i>Input Editor</i>	<i>Simple HCAD</i>	<i>Simple Planner</i>	<i>Low Fidelity Sim.</i>	<i>Output Visual.</i>
GIS	X				X
Statistical Analysis					X
Animation					X
Plotting					X
Database	X	X	X	X	X

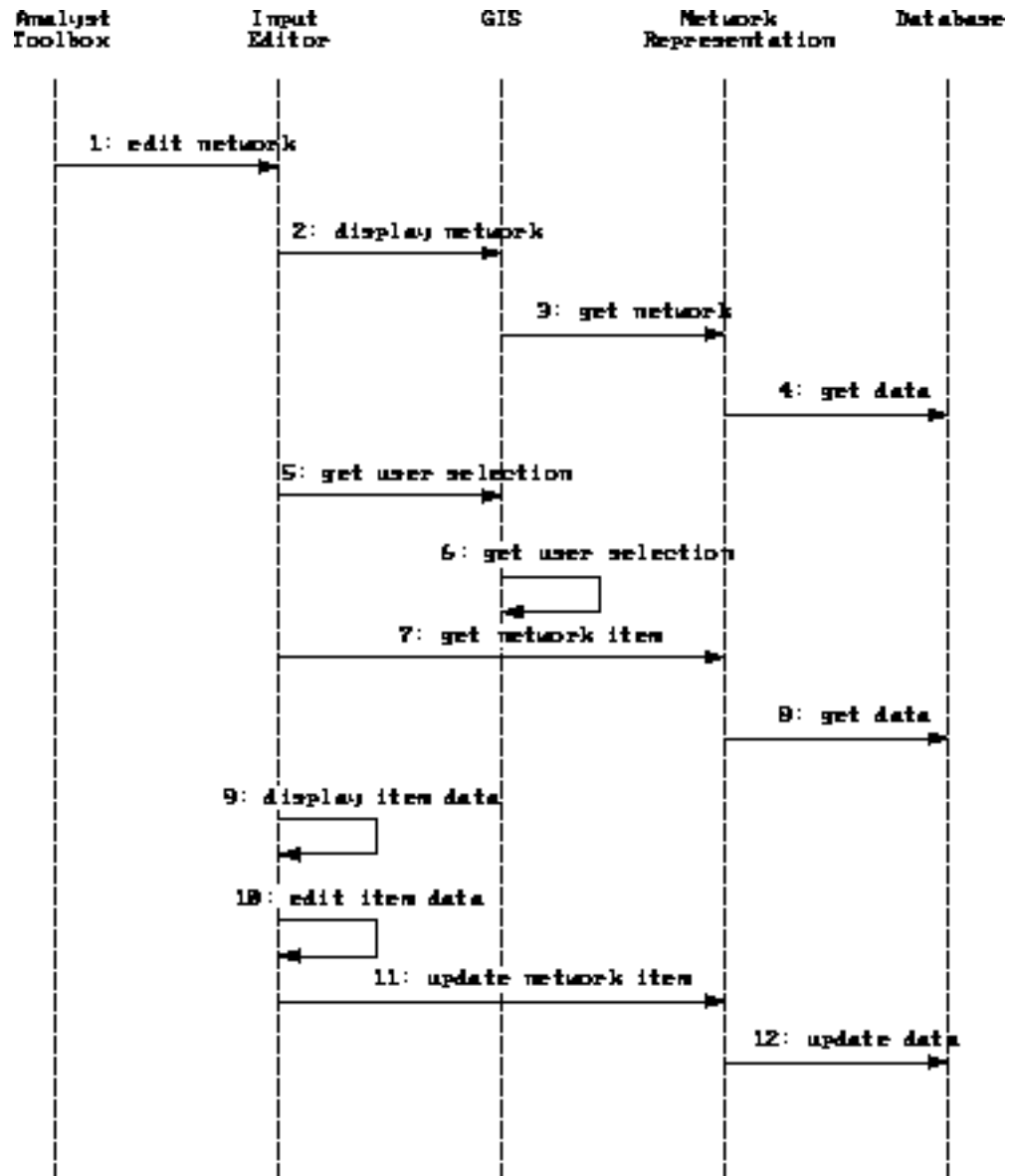
Data Flow Specifications

Subsystem Interaction Diagrams

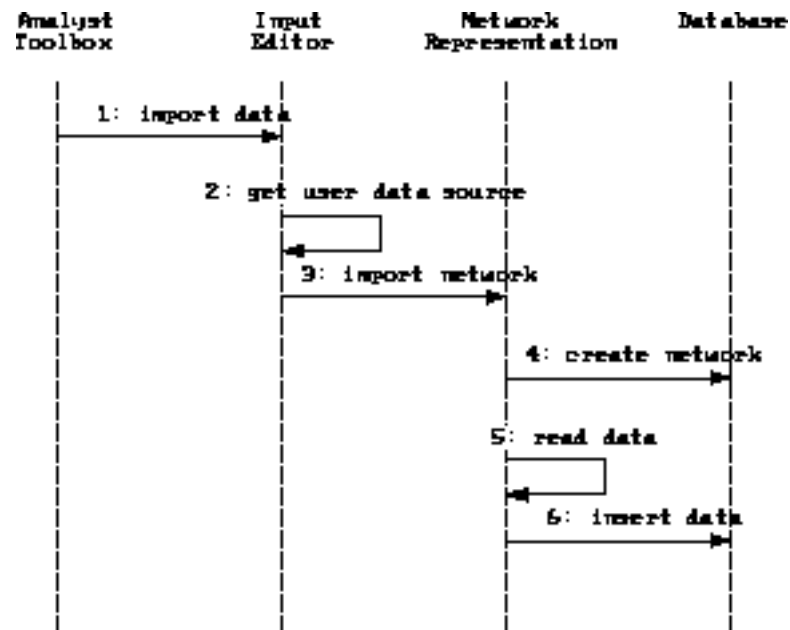
II. Data Flow Specifications.

A. Subsystem Interaction Diagrams.

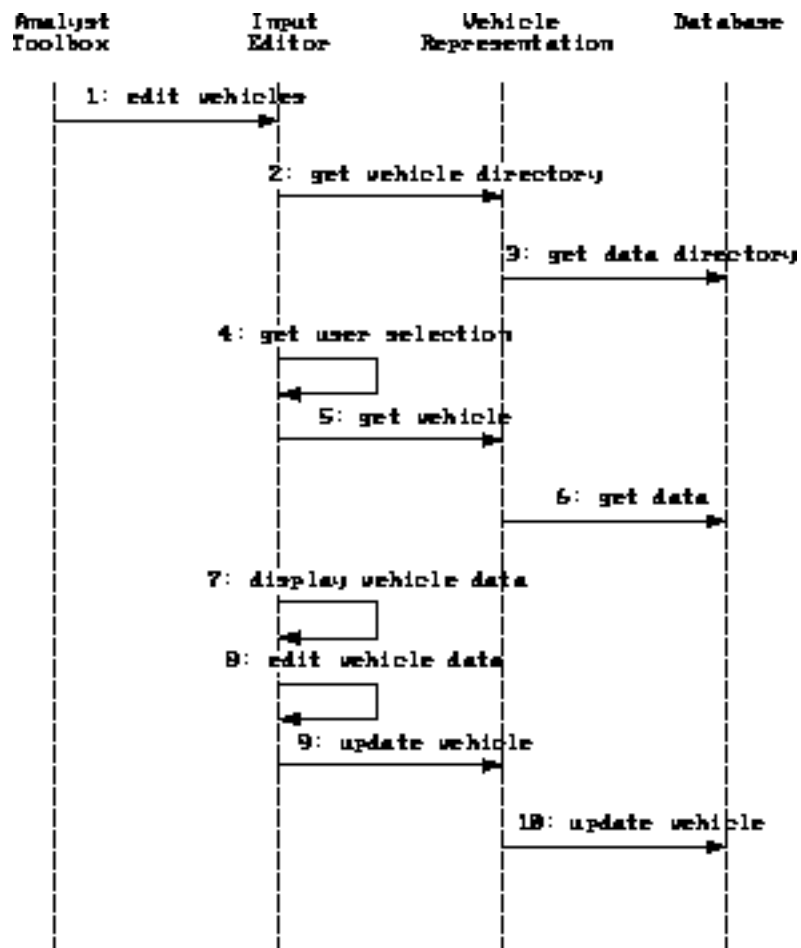
1. Edit network item.



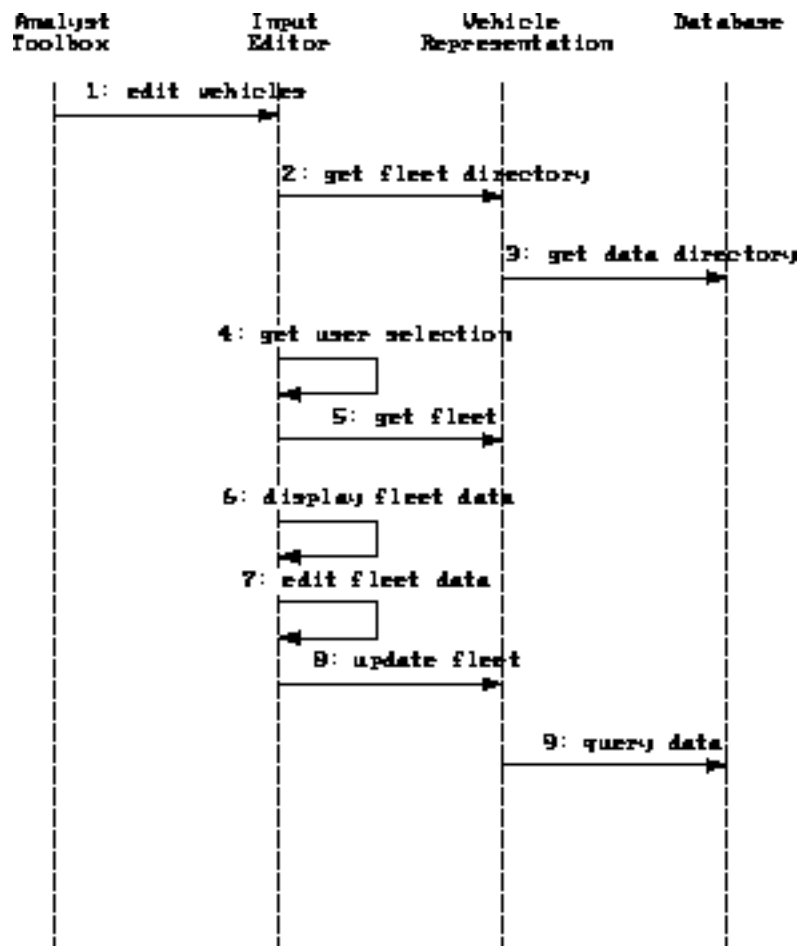
2. Import road network.



3. Edit vehicle data.

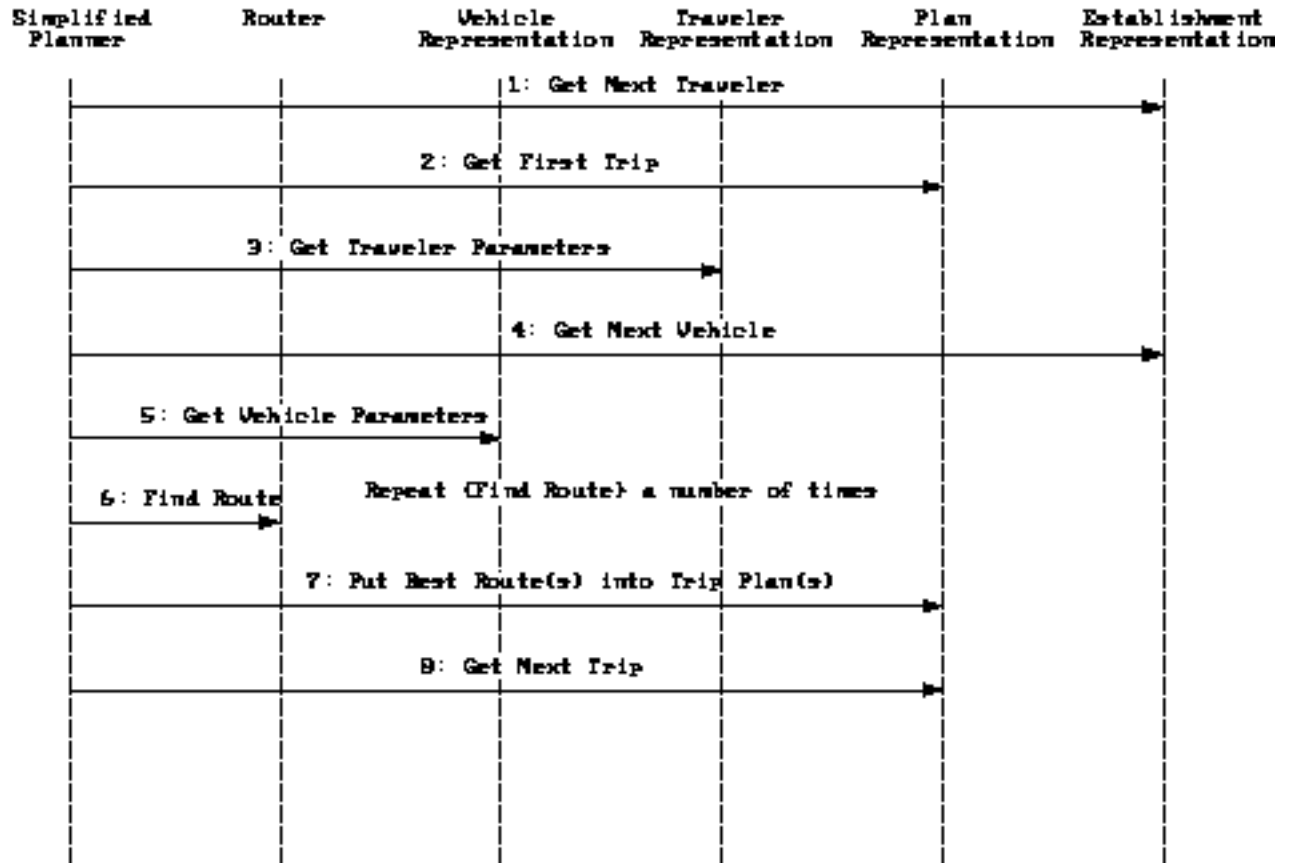


4. Edit fleet data.

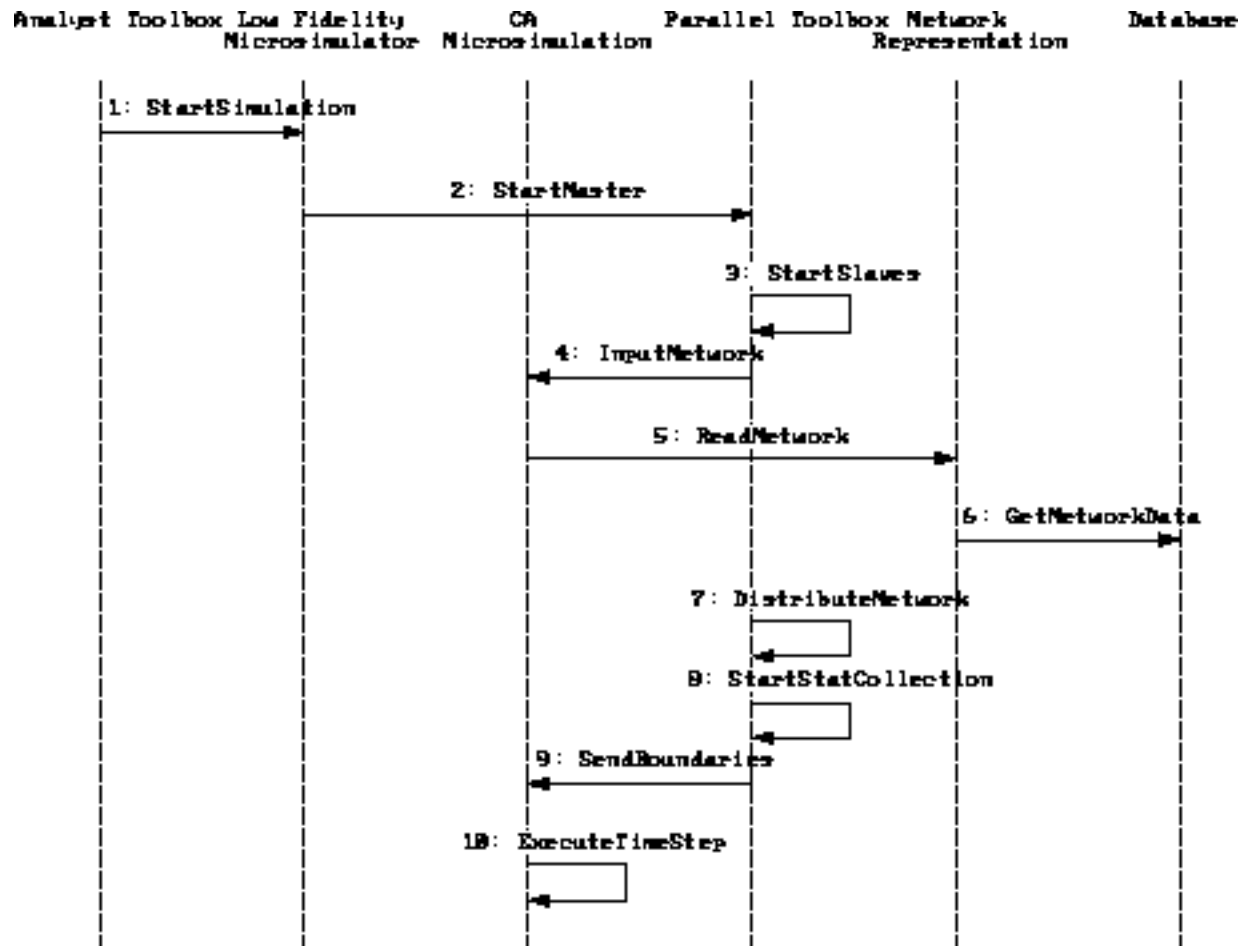


5. Make plan.

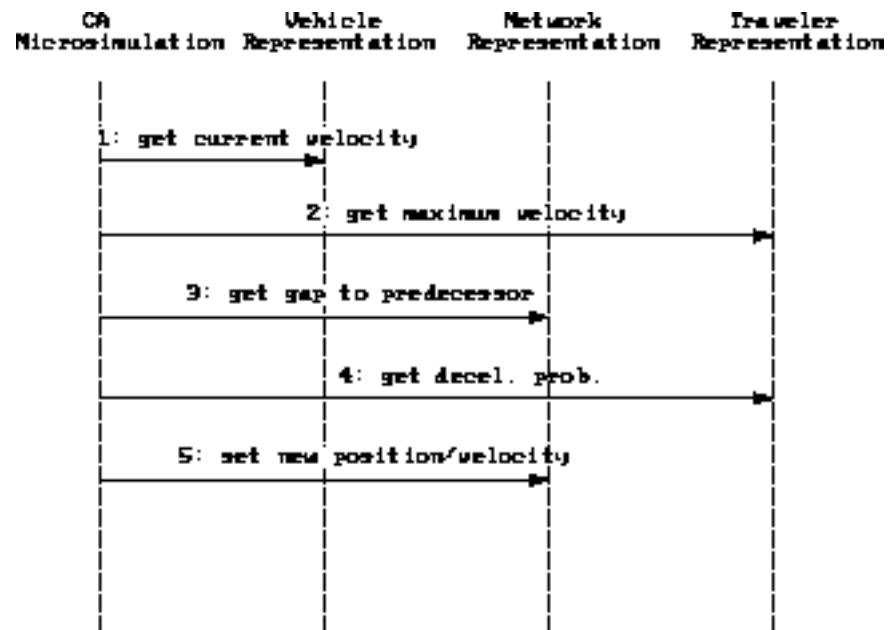
Simplified Planner gets the next traveler and finds the best route(s) for that traveler's first



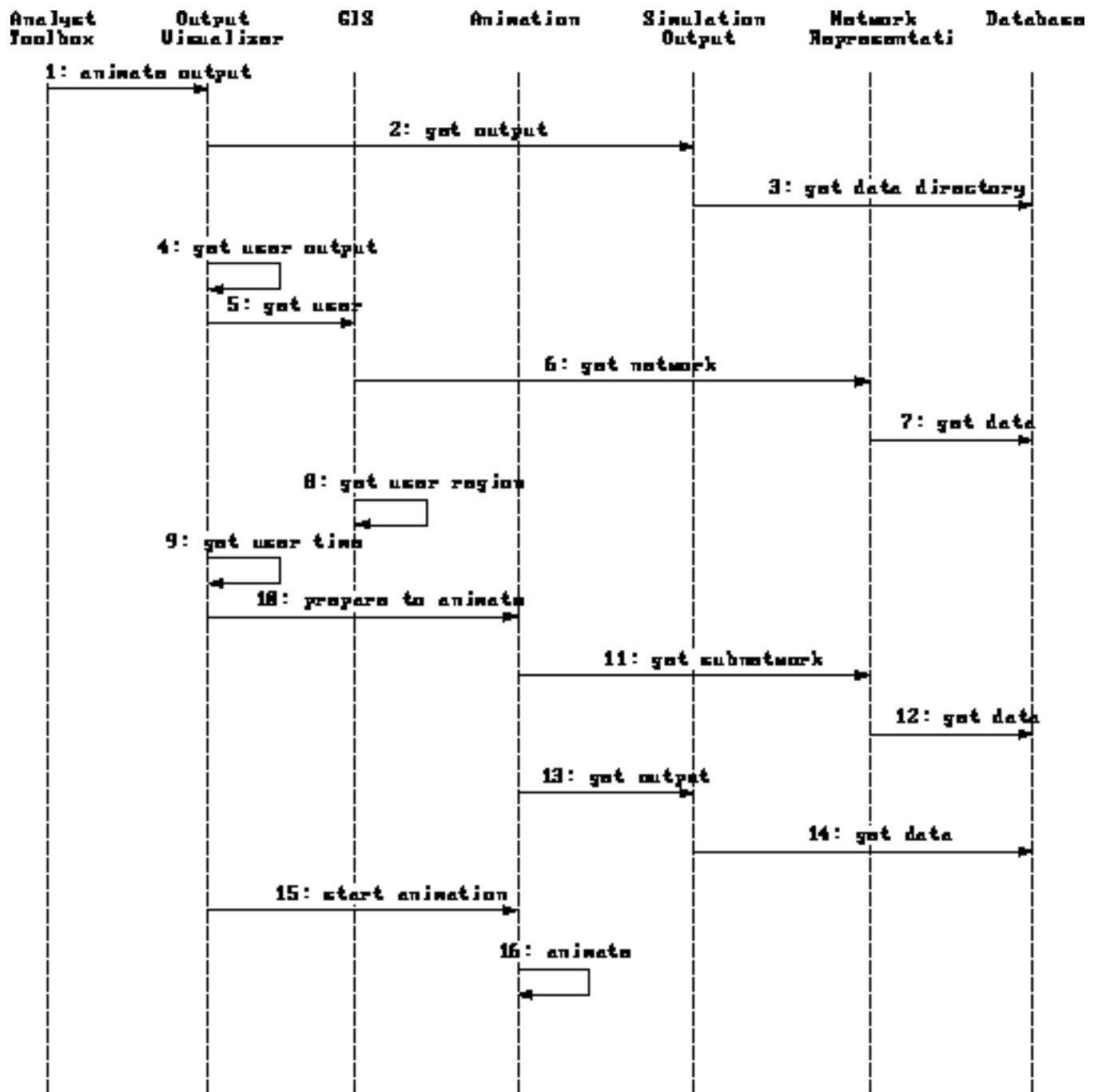
6. Start simulation.



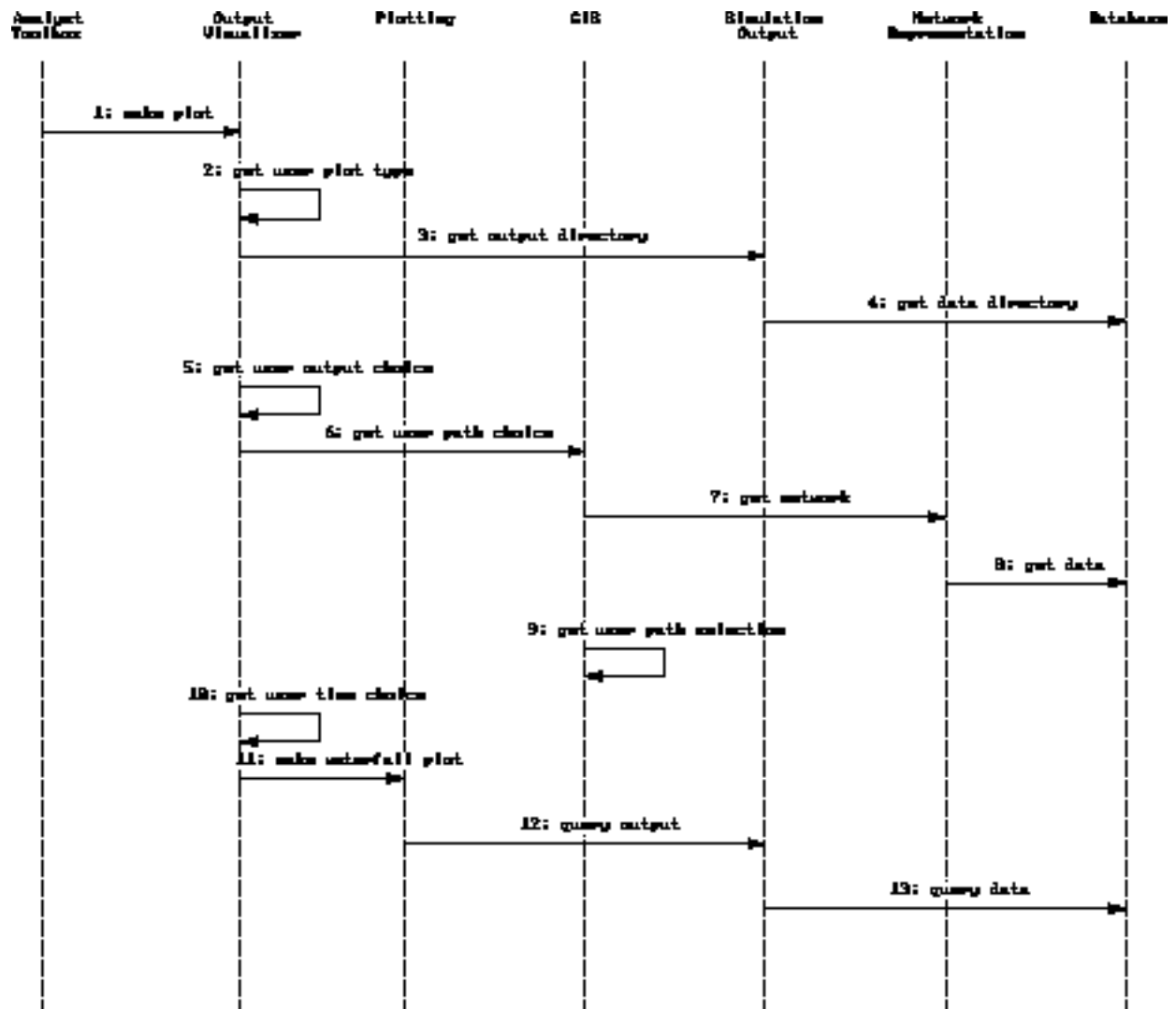
7. Move vehicle on link.



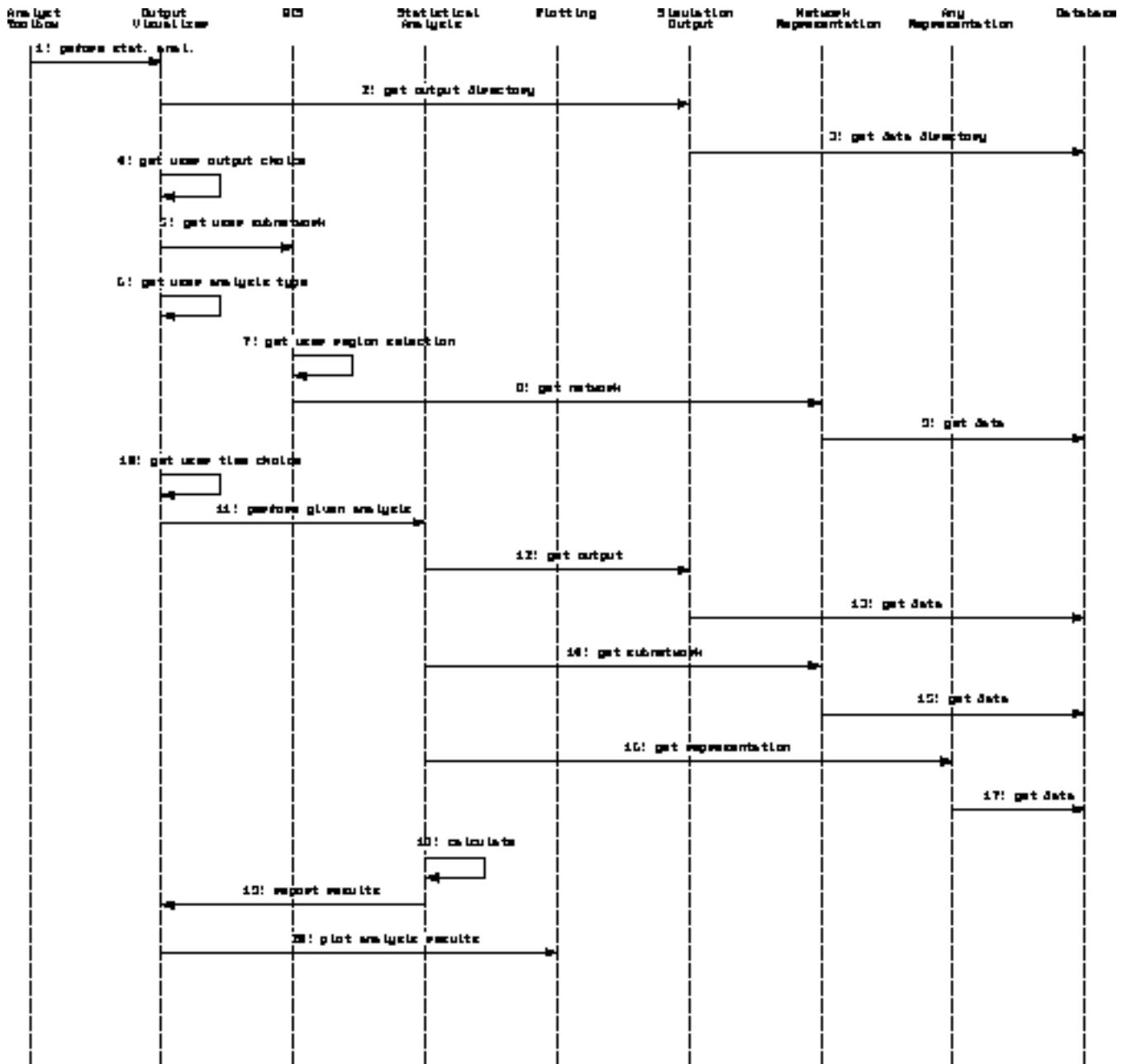
8 Animate simulation output.



9. Make waterfall plot.



10. Perform statistical analysis.



Entity-Relationship Diagram

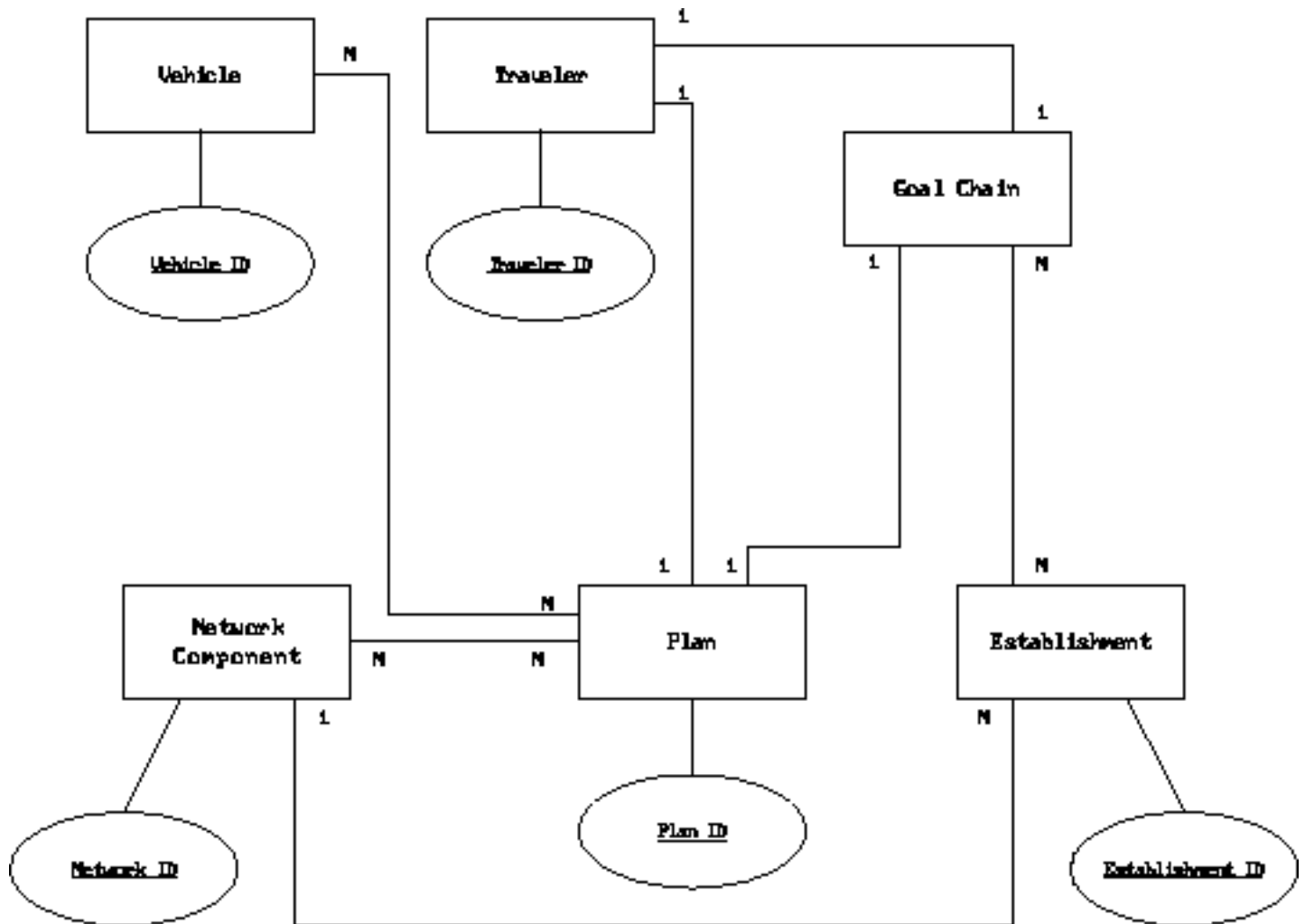
Each of the data items in a representation subsystem has a unique identifier (a 32 bit number).

This identifier may encode type information, as when a traveler identifier distinguishes between intra- and inter-region travelers, or between cargo and persons.

The identifier values are not unique between subsystems or for different simulation scenarios.

B. Entity-relationship diagram.

Each of the data items in a representation subsystem has a unique identifier (a 32 bit number). This identifier may encode type information, as when a traveler identifier distinguishes between intra- and inter-region travelers, or between cargo and persons. The identifier values are not unique between subsystems or for different simulation scenarios.



Review of Previous Prototypes

CA Microsimulator

This section summarizes our evaluation of the existing CA microsimulation prototype. Details of the analysis are provided in Attachment 1.

The existing CA microsimulation implements some of the capabilities required for IOC-1, but lacks others including multiple-lane bidirectional traffic, intersections, and drivers with plans and suitably complex driving behavior.

Additionally, load balancing among the distributed processors does not presently work for complex networks.

Nevertheless, we recommend that the existing code be enhanced to provide the required missing functionality.

This recommendation carries significant risk (addressed more below), but the alternative of starting over appears to be more risky given the desired late summer completion date for IOC-1.

The existing implementation is undocumented and therefore difficult to understand without extensive study.

At this time, we are unable to estimate how long it would take to reimplement all of it.

By retaining and enhancing the existing implementation, we can temporarily treat portions of the code as black boxes, to be examined as we discover problems or need to add functionality.

Considerable work has gone into the implementation of the parallel toolbox, and some effort has been made to logically separate the toolbox from the traffic simulation details.

We recommend continuing efforts to reduce the coupling between subsystems used by the microsimulation by developing interface classes where appropriate.

Documentation should be incorporated as classes are understood, and coding practices that enhance the discovery of problems at compile time should be introduced as classes are modified.

We also recommend migrating to the common set of container classes adopted for use in TRANSIMS (i.e., the Booch Components) as soon as practicable.

Finally, we recommend ongoing tracking of Rickert's independent enhancements to the current implementation and consideration for whether incorporating them into our version would be beneficial and cost-effective.

Graphical User Interface

This section summarizes our evaluation of the existing graphical user interface prototype. Details of the analysis are provided in Attachment 2.

The prototype graphical user interface was developed for TRANSIMS data visualization.

It was never intended as a toolbox to access the different TRANSIMS systems. Therefore, it is limited in its functionality.

The main TRANSIMS interface, the Analyst Toolbox, will be implemented using ArcView.

The animation of vehicular traffic is still necessary aside from what ArcView can provide.

Based on our detailed evaluation of the existing prototype, it is recommended that the prototype not be used for the animation subsystem in TRANSIMS.

An improved animation capability should be implemented.

C++ Planner

This section summarizes our evaluation of the existing C++ planner. Details of the analysis are provided in Attachment 3.

The I/O interfaces will have to be replaced, so the only reusability question concerns the class structure and the two key functions of finding routes and evaluating them.

For example, the router unwinds dead ends, prunes loops, and avoids circular routes. Its search strategy changes as it nears its destination. The Traveler and Plan classes are tightly coupled to it.

We recommend that these algorithms be evaluated for possible reuse.

The cost model weighs both geographic and demographic factors. However, its code module is easily modifiable or replaceable.

The source code needs some additional commenting for clarity. Several long functions (more than 100 lines) could be broken up to enhance maintainability.

At least one layer of call structure and at least one unused data member could be deleted.

Assessment of Major Risks

Cellular Automata

There may be unforeseen difficulties in evolving the present CA software into production code.

This could make certain enhancements of the CA very difficult, awkward, and time consuming.

Requirements

If the requirements for IOC-1 are not managed carefully, additions and changes may sidetrack the overall development effort.

A global prioritizing of the requirements is necessary.

Communication

If adequate communication is not maintained within the microsimulation team and between the microsimulation team and the other teams, software development work may be duplicated or discarded.

The interfaces between subsystems should be carefully communicated and reviewed.

Output

Without research and experimentation in the area of very-large database (VLDB) issues, it will be extremely difficult to manage the microsimulation data generated by a city-wide simulation.

The microsimulation team does not have the resources to address this issue in IOC-1.

Level of Effort Estimate

The table below indicates the relative level of effort (in arbitrary linear units) estimated to be required for the IOC-1 software development.

Each system or subsystem is an individual task that may be assigned to one or more software developers.

There are a total of 52 units of effort, NOT including the effort required to support a case study

Level of Effort Options

- (a) Not optimizing load balancing will reduce Parallel Toolbox by one unit.
- (b) Supporting only basic animation will reduce Animation by one unit.
- (c) Not supporting actuated signals will reduce Network Representation and CA Microsimulation by one unit each.
- (d) Doing a simulation of a large region (i.e., a whole city) will increase Simulation Output and Database by two units each.